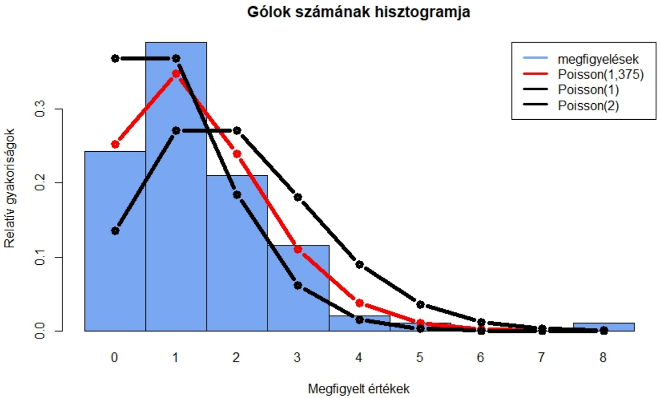
Poisson-eloszlás paraméterének becslése (4. előadás)



A gólok számának hisztogramja $n = 95$ mérkőzésen, és különböző paraméterű Poisson-eloszlások ($\mathbb{P}_\lambda(X = k) = \lambda^k/k! \cdot e^{-\lambda}$)

Poisson-eloszlás paraméterének becslése

Tegyük fel, hogy $X_1, X_2, \ldots, X_n$ független, azonos λ paraméterű Poisson-eloszlású minta, ahol $\lambda > 0$ ismeretlen paraméter, $n = 95$, és $\overline{X} = 1,379$.

Poisson-eloszlásnál:

$$\mathbb{P}_\lambda(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda}; \qquad \mathbb{E}(X_j) = \lambda; \qquad D(X_j) = \sqrt{\lambda}.$$

A megfigyelések az alábbiak (a gólok száma összesen $\sum_{j=1}^{n} X_j = 131$):

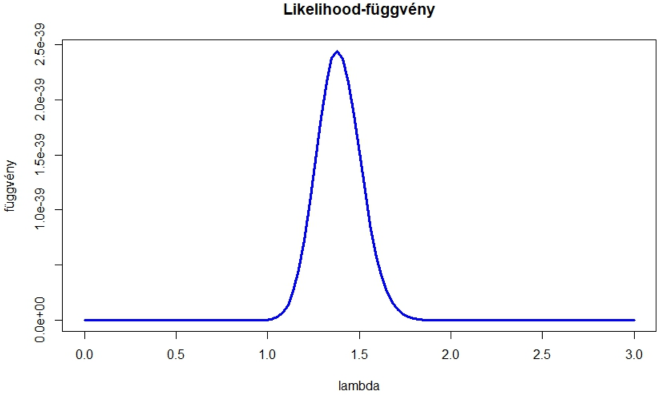
$$3, \quad 0, \quad 2, \quad 2, \quad 1, \quad 3, \ldots, 2.$$

Annak valószínűsége λ paraméter mellett, hogy éppen ezt a sorozatot kaptuk:

$$L_{95,\lambda}(3, 0, 2, \ldots, 2) = \frac{\lambda^3}{3!} e^{-\lambda} \cdot \frac{\lambda^0}{0!} e^{-\lambda} \cdot \frac{\lambda^2}{2!} e^{-\lambda} \cdot \ldots \cdot \frac{\lambda^2}{2!} e^{-\lambda} =$$

$$= \frac{\lambda^{3+0+2+\ldots+3}}{3! \cdot 0! \cdot 2! \cdot \ldots \cdot 2!} e^{-95\cdot\lambda} = \frac{\lambda^{131}}{3! \cdot 0! \cdot 2! \cdot \ldots \cdot 2!} e^{-95\cdot\lambda}$$

Poisson-eloszlás paraméterének becslése



A $\frac{\lambda^{131}}{3! \cdot 0! \cdot 2! \cdot \ldots \cdot 2!} e^{-95 \cdot \lambda}$ likelihoodfüggvény a $\lambda > 0$ paraméter függvényében;

mintaátlag: $\overline{X} = \frac{131}{95} = 1,379$

ML-becslés: Poisson-eloszlás

$X_1, \ldots, X_n$ függetlenek, Poisson-eloszlás $\lambda > 0$ ismeretlen paraméterrel, azaz

$$\mathbb{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, 2, \ldots).$$

Ekkor

$$L_{n,\lambda}(X_1, \ldots, X_n) = \prod_{j=1}^{n} \left(\frac{\lambda^{X_j}}{X_j!} e^{-\lambda} \right) = \frac{\lambda^{X_1}}{X_1!} e^{-\lambda} \cdot \frac{\lambda^{X_2}}{X_2!} e^{-\lambda} \cdot \ldots \cdot \frac{\lambda^{X_n}}{X_n!} e^{-\lambda}.$$

$$L_{n,\lambda}(X_1, \ldots, X_n) = \lambda^{\sum_{j=1}^{n} X_j} e^{-n\lambda} \cdot \prod_{j=1}^{n} \frac{1}{X_j!}.$$

ML-becslés: Poisson-eloszlás

$X_1, \ldots, X_n$ függetlenek, Poisson-eloszlás $\lambda > 0$ ismeretlen paraméterrel, azaz

$$\mathbb{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, 2, \ldots).$$

Ekkor

$$L_{n,\lambda}(X_1, \ldots, X_n) = \prod_{j=1}^{n} \left(\frac{\lambda^{X_j}}{X_j!} e^{-\lambda} \right) = \frac{\lambda^{X_1}}{X_1!} e^{-\lambda} \cdot \frac{\lambda^{X_2}}{X_2!} e^{-\lambda} \cdot \ldots \cdot \frac{\lambda^{X_n}}{X_n!} e^{-\lambda}.$$

$$L_{n,\lambda}(X_1, \ldots, X_n) = \lambda^{\sum_{j=1}^{n} X_j} e^{-n\lambda} \cdot \prod_{j=1}^{n} \frac{1}{X_j!}.$$

$$\log L_{n,\lambda}(X_1, \ldots, X_n) = \log \lambda \sum_{j=1}^{n} X_j - n\lambda - \log \prod_{j=1}^{n} \frac{1}{X_j!}$$

ML-becslés: Poisson-eloszlás

$X_1, \ldots, X_n$ függetlenek, Poisson-eloszlás $\lambda > 0$ ismeretlen paraméterrel, azaz

$$\mathbb{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (k = 0, 1, 2, \ldots).$$

Ekkor

$$L_{n,\lambda}(X_1, \ldots, X_n) = \prod_{j=1}^{n} \left(\frac{\lambda^{X_j}}{X_j!} e^{-\lambda} \right) = \frac{\lambda^{X_1}}{X_1!} e^{-\lambda} \cdot \frac{\lambda^{X_2}}{X_2!} e^{-\lambda} \cdot \ldots \cdot \frac{\lambda^{X_n}}{X_n!} e^{-\lambda}.$$

$$L_{n,\lambda}(X_1, \ldots, X_n) = \lambda^{\sum_{j=1}^{n} X_j} e^{-n\lambda} \cdot \prod_{j=1}^{n} \frac{1}{X_j!}.$$

$$\log L_{n,\lambda}(X_1, \ldots, X_n) = \log \lambda \sum_{j=1}^{n} X_j - n\lambda - \log \prod_{j=1}^{n} \frac{1}{X_j!}$$

$$\frac{\partial}{\partial \lambda} \log L_{n,\lambda}(X_1, \ldots, X_n) = \frac{\sum_{j=1}^{n} X_j}{\lambda} - n > 0 \Leftrightarrow \lambda < \overline{X}.$$

Ezért az ML-becslés: $\hat{\lambda} = \overline{X}$, mivel a derivált itt vált előjelet.

Maximumlikelihood-módszer

Definíció (Likelihood-függvény)

Ha az $(Y_1, \ldots, Y_n)$ *független minta diszkrét (a lehetséges értékeinek száma véges vagy megszámlálható sok), akkor a likelihood-függvénye:*

$$L_{n,\vartheta}(k_1, \ldots, k_n) = \prod_{j=1}^{n} \mathbb{P}_{j,\vartheta}(Y_j = k_j) \qquad ((k_1, \ldots, k_n) \in H).$$

Maximumlikelihood-módszer

Definíció (Likelihood-függvény)

Ha az $(Y_1, \ldots, Y_n)$ *független minta diszkrét (a lehetséges értékeinek száma véges vagy megszámlálható sok), akkor a likelihood-függvénye:*

$$L_{n,\vartheta}(k_1, \ldots, k_n) = \prod_{j=1}^{n} \mathbb{P}_{j,\vartheta}(Y_j = k_j) \qquad ((k_1, \ldots, k_n) \in H).$$

Ha az $(Y_1, \ldots, Y_n)$ *független minta abszolút folytonos, és* Y_j *sűrűségfüggvénye (a* $\mathbb{P}_\vartheta$ *valószínűség mellett)* $f_{j,\vartheta}$*, akkor a minta likelihood-függvénye:*

$$L_{n,\vartheta}(t_1, \ldots, t_n) = \prod_{j=1}^{n} f_{j,\vartheta}(t_j) \qquad (t_1, \ldots, t_n \in \mathbb{R}).$$

Maximumlikelihood-módszer

Legyen $(\Omega, \mathcal{A}, \mathbb{P})$ statisztikai mező, ahol $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$, vagyis az ismeretlen eloszlás a ϑ paraméterrel jellemezhető.

Definíció (Maximum-likelihood becslés)

A ϑ maximumlikelihood-becslése (ML-becslése) az $X_1, \ldots, X_n$ mintából $\hat{\vartheta}$, ha $\hat{\vartheta}$ maximalizálja a $\vartheta \mapsto L_{n,\vartheta}(X_1, \ldots, X_n)$ függvényt, ahol $L_{n,\vartheta}$ a minta likelihood-függvénye. Azaz, ha

$$L_{n,\hat{\vartheta}}(X_1, \ldots, X_n) \geq L_{n,\vartheta}(X_1, \ldots, X_n) \text{ minden } \vartheta \in \Theta\text{-ra}.$$

ML-becslés: normális eloszlás

$X_1, \ldots, X_n$ függetlenek, eloszlásuk normális eloszlás $m, \sigma > 0$ paraméterekkel. Ekkor

$$L_{n,m,\sigma}(X_1, \ldots, X_n) = \prod_{j=1}^{n} f_{j,\vartheta}(X_j) = \prod_{j=1}^{n} \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X_j - m)^2}{2\sigma^2} \right) \right].$$

ML-becslés: normális eloszlás

$X_1, \ldots, X_n$ függetlenek, eloszlásuk normális eloszlás $m, \sigma > 0$ paraméterekkel. Ekkor

$$L_{n,m,\sigma}(X_1, \ldots, X_n) = \prod_{j=1}^{n} f_{j,\vartheta}(X_j) = \prod_{j=1}^{n} \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(- \frac{(X_j - m)^2}{2\sigma^2} \right) \right].$$

$$L_{n,m,\sigma}(X_1, \ldots, X_n) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp\left(- \sum_{j=1}^{n} \frac{(X_j - m)^2}{2\sigma^2} \right).$$

ML-becslés: normális eloszlás

$X_1, \ldots, X_n$ függetlenek, eloszlásuk normális eloszlás $m, \sigma > 0$ paraméterekkel. Ekkor

$$L_{n,m,\sigma}(X_1, \ldots, X_n) = \prod_{j=1}^{n} f_{j,\vartheta}(X_j) = \prod_{j=1}^{n} \left[\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X_j - m)^2}{2\sigma^2} \right) \right].$$

$$L_{n,m,\sigma}(X_1, \ldots, X_n) = \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp\left(-\sum_{j=1}^{n} \frac{(X_j - m)^2}{2\sigma^2} \right).$$

$$\log L_{n,m,\sigma}(X_1, \ldots, X_n) = -n \log(\sqrt{2\pi}) - n \log \sigma - \sum_{j=1}^{n} \frac{(X_j - m)^2}{2\sigma^2}.$$

Rögzített σ mellett ez akkor maximális, ha $\sum_{j=1}^{n}(X_j - m)^2 = \sum_{j=1}^{n} X_j^2 - 2\sum_{j=1}^{n} X_j m + nm^2$ minimális $\Rightarrow \hat{m} = \overline{X}$.

ML-becslés: normális eloszlás

$$\log L_{n,\sigma}(X_1, \dots, X_n) = -n\log(\sqrt{2\pi}) - n\log\sigma - \sum_{j=1}^{n} \frac{(X_j - \overline{X})^2}{2\sigma^2}.$$

A σ szerinti parciális derivált:

$$\frac{\partial}{d\sigma} \log L_{n,\sigma}(X_1, \dots, X_n) = -\frac{n}{\sigma} + \sum_{j=1}^{n} \frac{(X_j - \overline{X})^2}{\sigma^3}.$$

Ez pontosan akkor pozitív, ha $\sigma^2 < \frac{1}{n}\sum_{j=1}^{n}(X_j - \overline{X})^2 = s_n^2$.

Tehát az ML-becslés:

$$\hat{m} = \overline{X}; \qquad \hat{\sigma} = s_n = \sqrt{\frac{1}{n}\sum_{j=1}^{n} X_j^2 - \overline{X}^2}.$$

Tehát normális eloszlásnál az m paraméter becslése a mintaátlag, a szórásé a tapasztalati szórás.

Az ML-becslés tulajdonságai

- Nem minden statisztikai mezőn létezik ML-becslés.
- Az ML-becslés nem feltétlenül egyértelmű.
- Az ML-becslés nem feltétlenül torzítatlan.
- A $\psi(\vartheta)$ függvény ML-becslése $\psi(\hat{\vartheta})$, ahol $\hat{\vartheta}$ ML-becslés ϑ-ra.
- **Megfelelő feltételek** (erős regularitási feltételek mellett) az ML-becslés aszimpotikusan torzítatlan, **aszimptotikusan hatásos** (vagyis minimális szórású) és **aszimptotikusan normális eloszlású**, azaz $\sqrt{n}(\hat{\vartheta}_n - \vartheta)$ normális eloszláshoz konvergál eloszlásban $n \to \infty$ esetén (a $\mathbb{P}_\vartheta$ valószínűségre vonatkozóan).
- Az alábbi egyenlet a maximumlikelihood-egyenlet:

$$\frac{\partial}{\partial\vartheta} \log L_{n,\vartheta}(X_1, \ldots, X_n) = 0.$$

Megfelelő feltételek mellett az ML-becslés a maximumlikelihood-egyenlet megoldása (ha az ML-becslés nem számítható ki, de az egyenlet megoldható, gyakran az egyenlet megoldásával helyettesítik az ML-becslést).

ML-becslés: egyenletes eloszlás

Ebben az esetben az ML-becslés nem számítható ki az ML-egyenlet gyökeként, vagyis nem kapható meg deriválással.

$X_1, \ldots, X_n$ függetlenek, eloszlásuk egyenletes eloszlás az $[a, b]$ intervallumon. Ekkor

$$L_{n,a,b}(X_1, \ldots, X_n) = \prod_{j=1}^{n} f_{j,\vartheta}(X_j) = \prod_{j=1}^{n} \mathbb{I}(a \leq X_j \leq b) \cdot \frac{1}{b-a}.$$

$$L_{n,a,b}(X_1, \ldots, X_n) = \left(\frac{1}{b-a}\right)^n \mathbb{I}(a \leq \min_j X_j \text{ és } \max_j X_j \leq b).$$

ML-becslés: egyenletes eloszlás

Ebben az esetben az ML-becslés nem számítható ki az ML-egyenlet gyökeként, vagyis nem kapható meg deriválással.

$X_1, \ldots, X_n$ függetlenek, eloszlásuk egyenletes eloszlás az $[a, b]$ intervallumon. Ekkor

$$L_{n,a,b}(X_1, \ldots, X_n) = \prod_{j=1}^{n} f_{j,\vartheta}(X_j) = \prod_{j=1}^{n} \mathbb{I}(a \leq X_j \leq b) \cdot \frac{1}{b-a}.$$

$$L_{n,a,b}(X_1, \ldots, X_n) = \left(\frac{1}{b-a}\right)^n \mathbb{I}(a \leq \min_j X_j \text{ és } \max_j X_j \leq b).$$

Az első tényező legyen minél nagyobb (vagyis $b - a$ minél kisebb) úgy, hogy a második tényező nem nulla. Ebből:

$$\hat{a} = \min_j X_j; \qquad \hat{b} = \max_j X_j.$$

Maximum-likelihood becslések

- binomiális eloszlás ismert k renddel: $\hat{p} = \overline{X}/k$
- Poisson-eloszlás: $\hat{\lambda} = \overline{X}$
- geometriai eloszlás: $\hat{p} = 1/\overline{X}$
- normális eloszlás: $\hat{m} = \overline{X}, \hat{\sigma} = s_n$
- exponenciális eloszlás: $\hat{\lambda} = 1/\overline{X}$
- egyenletes eloszlás: $\hat{a} = \min_j X_j; \qquad \hat{b} = \max_j X_j$

Momentummódszer ismeretlen paraméterek becslésére

Legyen $X_1, \dots, X_n$ független azonos eloszlású minta.

1. Az eloszlás k. momentuma, ha ϑ az ismeretlen paraméter: $\mu_{k,\vartheta} = \mathbb{E}_\vartheta(X_1^k)$.
2. Legyen $\hat{\mu}_k = \frac{1}{n}\sum_{j=1}^n X_j^k$ az eloszlás k. tapasztalati momentuma.
3. Írjuk fel az alábbi egyenleteket a legkisebb olyan k-ig, amire az egyenletrendszer egyértelműen meghatározza ϑ-t (**bár nincs mindig ilyen k**):

$$\mathbb{E}_\vartheta(X_1) = \frac{1}{n}\sum_{j=1}^n X_j;$$

$$\mathbb{E}_\vartheta(X_1^2) = \frac{1}{n}\sum_{j=1}^n X_j^2;$$

$$\dots$$

$$\mathbb{E}_\vartheta(X_1^k) = \frac{1}{n}\sum_{j=1}^n X_j^k.$$

4. A ϑ momentummódszerrel kapott becslése az a $\hat{\vartheta}$, ami megoldása a fenti egyenletrendszernek. **Nem mindig létezik, nem mindig egyértelmű, nem feltétlenül hatásos.**

Momentummódszer: Poisson– és exponenciális eloszlás

$X_1, \dots, X_n$ független **Poisson-eloszlásúak** ismeretlen $\lambda > 0$ paraméterrel. A $k = 1$-hez tartozó egyenlet:

$$\mathbb{E}_\lambda(X_1) = \overline{X}.$$

Mivel a λ paraméterű Poisson-eloszlás várható értéke λ:

$$\hat{\lambda} = \overline{X}.$$

Momentummódszer: Poisson– és exponenciális eloszlás

$X_1, \ldots, X_n$ független **Poisson-eloszlásúak** ismeretlen $\lambda > 0$ paraméterrel. A $k = 1$-hez tartozó egyenlet:

$$\mathbb{E}_\lambda(X_1) = \overline{X}.$$

Mivel a λ paraméterű Poisson-eloszlás várható értéke λ:

$$\hat{\lambda} = \overline{X}.$$

$X_1, \ldots, X_n$ független **exponenciális** eloszlásúak ismeretlen $\lambda > 0$ paraméterrel. A $k = 1$-hez tartozó egyenlet:

$$\mathbb{E}_\lambda(X_1) = \frac{1}{\lambda} = \frac{1}{n}\sum_{j=1}^{n} X_j = \overline{X}.$$

Ez egyértelműen oldható meg λ-ra:

$$\hat{\lambda} = \frac{1}{\overline{X}}.$$

Momentummódszer: normális eloszlás

$X_1, \ldots, X_n$ független $N(m, \sigma^2)$ eloszlású minta (azaz normális eloszlású m várható értékkel és σ szórással).

A $k = 1$-hez és $k = 2$-höz tartozó egyenletek:

$$\mathbb{E}_{m,\sigma}(X_1) = m = \overline{X};$$
$$\mathbb{E}_{m,\sigma}(X_1^2) = \sigma^2 + m^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2.$$

A másodikba beírva az elsőt: $\sigma^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \overline{X}^2 = s_n^2$ (a tapasztalati szórásnégyzet). Tehát az első két egyenlet együtt egyértelműen oldható meg, a momentummódszerrel kapott becslés:

$$\hat{m} = \overline{X}; \qquad \hat{\sigma} = s_n.$$

Momentummódszer: egyenletes eloszlás

Legyen $X_1, \ldots, X_n$ független minta az $[a, b]$ intervallumon egyenletes eloszlásból. Ennek várható értéke $(a+b)/2$, szórása $(b-a)/\sqrt{12}$. Ezek alapján a $k=1$-hez és $k=2$-höz tartozó egyenlet:

$$\mathbb{E}_{a,b}(X_1) = \frac{a+b}{2} = \overline{X};$$

$$\mathbb{E}_{a,b}(X_1^2) = \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2.$$

A másodikba beírva az elsőt: $\frac{(b-a)^2}{12} = \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \overline{X}^2 = s_n^2$, amiből

$$\hat{a} = \overline{X} - \sqrt{3}s_n; \qquad \hat{b} = \overline{X} + \sqrt{3}s_n.$$

Elégséges statisztika

Definíció

Legyen $X_1, \ldots, X_n$ *független minta* ϑ *paraméterű eloszlásból. A* T *statisztika elégséges, ha a likelihood-függvény felírható a következő alakban megfelelő* h *és* g *függvényekkel:*

$$L_{n,\vartheta}(X_1, \ldots, X_n) = h(X_1, \ldots, X_n) \cdot g_\vartheta(T(X_1, \ldots, X_n)).$$

Elégséges statisztika

Definíció

Legyen $X_1, \dots, X_n$ független minta ϑ paraméterű eloszlásból. A T statisztika elégséges, ha a likelihood-függvény felírható a következő alakban megfelelő h és g függvényekkel:

$$L_{n,\vartheta}(X_1, \dots, X_n) = h(X_1, \dots, X_n) \cdot g_\vartheta(T(X_1, \dots, X_n)).$$

Például a Poisson-eloszlás esetén $T(X_1, \dots, X_n) = \overline{X}$, normális eloszlás esetén $T(X_1, \dots, X_n) = (\overline{X}, s_n^2)$ elégséges statisztika.

Az elégséges statisztika nem egyértelmű.

Ha létezik ML-becslés, T pedig elégséges statisztika, akkor az ML-becslés felírható $s(T(X_1, \dots, X_n))$ alakban valamely s függvényre.

Rao–Blackwell-tétel

Tétel (Rao–Blackwell)

Tegyük fel, hogy a T statisztika torzítatlan becslés a ϑ paraméterre, valamint S elégséges statisztika ϑ-ra. Ekkor megadható olyan T^* becslés, melyre

- $T^* = h(S)$ *megfelelő h függvénnyel;*
- T^* *torzítatlan ϑ-ra: $\mathbb{E}_\vartheta(T^*) = \vartheta$ minden $\vartheta \in \Theta$-ra;*
- T^* *hatásosabb, mint T: $D_\vartheta(T^*) \le D_\vartheta(T)$ minden $\vartheta \in \Theta$-ra.*

Ha az S statisztika teljes is (azaz minden olyan f függvény, melyre $\mathbb{E}_\vartheta(f(S)) = 0$ minden ϑ-ra, „majdnem mindenütt" nulla), akkor T^* hatásos, azaz minden torzítatlan becslésnél hatásosabb.

A megoldás az $\mathbb{E}(T|S)$ feltételes várható érték lesz.

Állítás

Poisson, exponenciális, normális eloszlásoknál (m-re) a mintaelemek összege teljes elégséges statisztika, a mintaátlag pedig hatásos becslése a paraméternek.

Fisher-információ

Egy mintaelem **Fisher-információja**:

$$I_1(\vartheta) = \mathbb{E}_\vartheta\left(\left(\frac{\partial}{\partial\vartheta}\log f_\vartheta(X_1)\right)^2\right),$$

ahol f_ϑ a sűrűségfüggvény $\mathbb{P}_\vartheta$ mellett.

Fisher-információ

Egy mintaelem **Fisher-információja**:

$$I_1(\vartheta) = \mathbb{E}_\vartheta\left(\left(\frac{\partial}{\partial\vartheta}\log f_\vartheta(X_1)\right)^2\right),$$

ahol f_ϑ a sűrűségfüggvény $\mathbb{P}_\vartheta$ mellett.

- Megfelelő (regularitási) feltételek mellett a független azonos eloszlású, n elemű minta Fisher-információja: $I_n(\vartheta) = n \cdot I_1(\vartheta)$.
- Ha T elégséges statisztika, akkor $T(X_1, \ldots, X_n)$ Fisher-információja ugyanaz, mint $(X_1, \ldots, X_n)$ Fisher-információja (például Poisson-eloszlásnál az átlag Fisher-információja ugyanaz, mint a teljes mintáé).
- **Cramér–Rao-egyenlőtlenség**: megfelelő (regularitási) feltételek mellett, ha T torzítatlan becslés ϑ-ra, akkor

$$D^2(T(X)) \geq \frac{1}{I_n(\vartheta)} = \frac{1}{nI_1(\vartheta)}.$$

(Ebben az értelemben feleakkora szóráshoz négyszer annyi mintaelem kell.)